Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 116 (2017) 351–357

Procedia Computer Science

www.elsevier.com/locate/procedia

2nd International Conference on Computer Science and Computational Intelligence 2017, ICCSCI 2017, 13-14 October 2017, Bali, Indonesia

COMPARISON BETWEEN NEURAL NETWORK AND SUPPORT VECTOR MACHINE IN OPTICAL CHARACTER RECOGNITION

Michael Reynaldo Phangtriastu[a],Jeklin Harefa[a]*,Dian Felita Tanoto[a]

[a]*Computer Science Department, School of Computer Science, Bina Nusantara University, Jl. KH Syahdan 9, Jakarta 11480, Indonesia*

Abstract

Optical Character Recognition is one of the popular area in artificial intelligence and pattern recognition area. Generally, this technique converts the input image into an editable format in computer. This paper uses several techniques as a comparison for some extracted features, such as: zoning algorithm, projection profile, Histogram of Oriented Gradients (HOG) and combination of those feature extractions (zoning + projection, projection + HOG, zoning + HOG, zoning + projection + HOG). For the evaluation of the proposed system, this paper compare the most commonly classifiers: Support Vector Machine (SVM) and Artificial Neural Network (ANN). This experiment achieves the highest accuracy of 94.43% using Support Vector Machine (SVM) classifier with the feature extraction algorithms are:projection profile and the combination of zoning + projection profile

© 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 2nd International Conference on Computer Science and Computational Intelligence 2017.

*Keywords:*Optical Character Recognition, Support Vector Machine, Artificial Neural Network

1. Introduction

Optical Character Recognition (OCR) is a technique that allows convertingthe printed text into an editable format in computer. OCR is one of popular research area in artificial intelligence and pattern recognition since it makes possible to read the text and convert it into electronic files, which can be edited. Many applications have been

* Corresponding author. Tel.: +6221-534-5830; fax: +6221-530-0244.
*E-mail address:*jharefa@binus.edu

1877-0509 © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 2nd International Conference on Computer Science and Computational Intelligence 2017.
10.1016/j.procs.2017.10.061

developed using OCR. Some of them are captcha images[1], automatic number plate recognition[2,3], text information extraction[4].

There are several problems that occur in developing an OCR system. Firstly, the computer might be difficult when recognizing value that has similarity with other value. For example: it is hard for computer can determine the letter 'l' and number '1' correctly. The next problem comes when extracting features that have background noise, such as the contrast of font and paper, the picture behind the font, and many more.

Many researchers have been developed in order to get the better accuracy in Optical Character Recognition (OCR). One of the researches comes from Malon and his team in 2008 with study entitled "Mathematical symbol recognition with Support Vector Machines". This study discussed about recognizing the mathematical symbols in the figure using the Support Vector Machine (SVM)[5]. Support Vector Machine is used to improve the performance of InftyReader method in classifying the mathematical symbols, and the SVM successfully reduces the error up to 41%.

In the next study, research on Character Recognition is conducted by Rao and his team in 2016, in a study entitled "Optical Character Recognition Technique Algorithms"[6]. This paper presents a modified back propagation based method for optical character recognition. The proposed method successfully computes error rate and gives promising accuracy of 100% in OCR.

In 2015, Mahto, Bahtia and Sharma[7] proposed a combination of horizontal and vertical projection feature extraction in recognizing the hand-written characters of Gurmukhi. This experiment use SVM (linear and polynomial kernel) and k-NN (k=1, k=3, k=5, k=7) for classifying the hand-written characters. The result was claimed to give the highest accuracy of 98.06%by using linear SVM.

Murthy and Hanmandlu[8] introduced a new technique using zoning based feature extraction for character recognition. In this paper, the character of black pixel location is also taken into account in order to get the unique features as much as possible. The performance shows above 98.5% accuracy by using SVM classifier.

Based on those previous studies, this paper focuses on taking the typewritten characters as the input, process it using several feature extraction algorithms and recognizes the character based on the extracted features. This paper aim at attempting to compare some of the common feature extraction in character recognition, which are projection histogram, zoning algorithm, Histogram of Oriented Gradients (HOG) and the combination of those feature extractions (zoning + projection, projection + HOG, zoning + HOG, zoning + projection + HOG). To the author's best knowledge, there is no work have been done in combining those methods mentioned to extract the image features for optical character recognition. The learning methods, using Artificial Neural Network (ANN) and Support Vector Machine will be applied as classifier in this paper.

2. Proposed Method

In this section, Optical Character Recognition using combination of feature extractions and classifiers is proposed. Generally, there are several steps for translating characters from image into a form that can be manipulated. The stages are including: Feature Extraction, Classification and Recognition. The preprocessing stage is not included in this experiment since the used dataset already had the high quality of images. Before the dataset get into classifier, the features must be extracted first. This experiment uses Zoning, Projection, and Histogram of Oriented Gradients (HOG) and the combination of those methods (zoning + projection, projection + HOG, zoning + HOG, projection + zoning + HOG). After that, for the evaluation of proposed system, the two classifiers: Support Vector Machine (SVM) and Neural Network (NN) are used to achieve the best accuracy. The schematic diagram of the proposed system is shown in Figure 1.

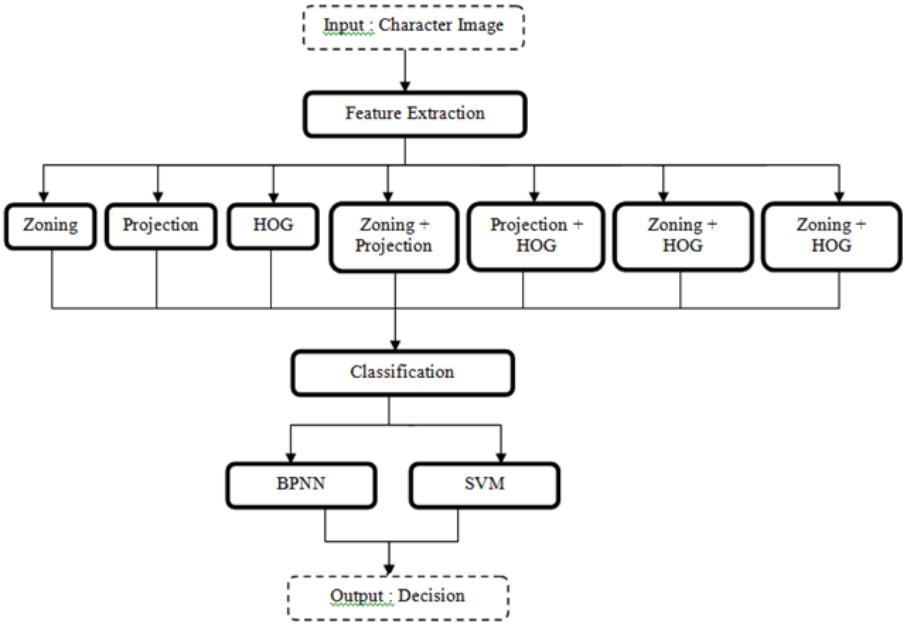

Figure 1.The Proposed System

3. Data Collection and Feature Extraction

3.1. Materials

This experiment uses English and Kannada characters which taken from the Chars74K dataset[9]. For English Characters, this dataset consists of 64 classes from 0-9, a-z and A-Z, 7705 characters of natural images, 3401 hand drawn characters using PC, 62992 synthesized characters from computer fonts. This dataset gives over 74K character images that explain the name of dataset. In Kannada Characters, the symbols use the combination of consonant and a vowel leads to a third class in this dataset.

This experiment uses English Fnt set which consists of 0-9, a-z and A-Z characters (black-white) from computer fonts with 4 variations (italic, bold, normal, and combination of bold and italic) and dimensions of 128 x 128 pixels. Each character has 1016 image with different styles. This experiment is limited to the capitalized character recognition with the composition of training and testing data set is 70:30. It means that there are 18491 images for training data and 7925 images of testing data and the total of the used images is 26416.Figure 1 shows the example of characters used in this experiment.

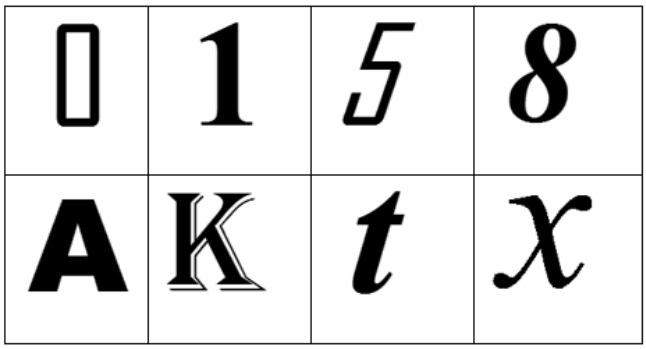

Figure 2. (a) Offline Character Recognition (b) Online Character Recognition

3.1.1. Zoning Algorithm

Zoning algorithm is one of popular technique for extracting features in pattern recognition area[10]. Zoning is a method for local information analysis on partitions of a given pattern. In this experiment, each of the character images are divided into 16zones (4 x 4) as in Figure 3. So, each zone contains 64 x 64 pixel. After the image already divided, the features that exist in each partsare computed(getthe average pixel density feature using equation (1):

$$d(i) = \frac{x}{n} \tag{1}$$

Where $d(i)$ is the density feature of zone-i , x is the number of foreground pixels, and n is the total number of pixels.

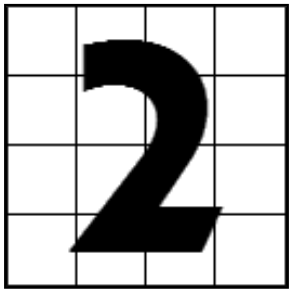

Figure 3. Zoning 4 x 4

In this experiment, there will be 16 density features that will be used as the input of classifier (as shown in Figure 3).

3.1.2. Projection Profile

Projection profile is one of feature extraction that accumulates the black pixel along rows and columns in the image. Basically, there are two types of projection profile, which are vertical and horizontal profiles. Let $S(N, M)$ denotes a binary image in N lines and M columns. For Vertical and Horizontal Profile, the equation can be shown in (2) and (3)[11].

$$P_v[i] = \sum_{j=1}^{M} S[i,j] \tag{2}$$

Where vertical profile (represented by the vector Pv of size M) is sum of the black pixel to the vertical axis.

$$P_h[j] = \sum_{i=1}^{N} S[i,j] \tag{3}$$

Where horizontal profile (represented by the vector Ph of size N) is sum of the black pixel to the vertical axis.
This experiment uses the combination of vertical and horizontal profile since this method already tested and achieves the accuracy of above 90%[7]. This researchuses 256 features which is obtained from 128 features of Horizontal Projection and 128 features of Vertical Projection.

3.1.3. Histogram of Oriented Gradients (HOG)

Histogram of Oriented Gradients (HOG) method is widely used in computer vision and image processing area for localization and detection of objects[12]. HOG operates on the gradient magnitude of the image. It is works by counting the occurrences of gradient orientations in localized parts of an image. Then, the image is divided into several cells and the histograms of gradient directions are formed for each of these cells (since each cell has a fixed gradient orientation bins). To achieve the better accuracy, the histograms have been normalized based on the contrast by calculating a measure of the intensity across block of images and using this value to normalize all cells within a block. This technique results a stable HOG against illumination and shadow variations[13].

In this experiment, the original size of each images are 128 x 128. Subsequently, every image will be divided into 64 x 64 format cells with the numbers of orientation bins are 9. This approach aim at producing 36 features that can be used to feed into the classifier.

4. OCR Classification

4.1.1. Support Vector Machine (SVM)

Support Vector Machine is a supervised learning method for classification and regression analysis. This algorithm output the optimal hyper-plane and maximizes the margin between two classes. For overcome the problem of non-linear data, SVM has the kernel trick that can obtain the better accuracy. There are several most common kernels for SVM. The simplest kernel is Linear Kernel[14]. The most favorite kernel is Gaussian Kernel[15]. And a kernel that suitable for the problem with normalized data is Polynomial kernel[16].This experiment uses the Gaussian kernel. It also called Radial Basis Function kernel (RBF kernel), which takes the following form:

$$K(x,x') = \exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) \tag{4}$$

where $\|x_i - x_j\|^2$ is recognized as square Euclidean distance between two vectors and σ is a kernel parameter.

4.1.2. Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is the algorithm which are designed to resemble the neural network brain system used for pattern recognition based on the input and output data[6]. It is also can be used to find the pattern of data or map the relationship between inputs and outputs.

The algorithm used in this experiment is Back Propagation Neural Network (BPNN) by reason of it is the simplest model to work and implements[17]. This algorithm generates an appropriate model that can be used to map the output based on the input data. There are three layers of BPNN structure: Input Layer, Hidden Layer and Output Layer as can be seen in Figure 4.

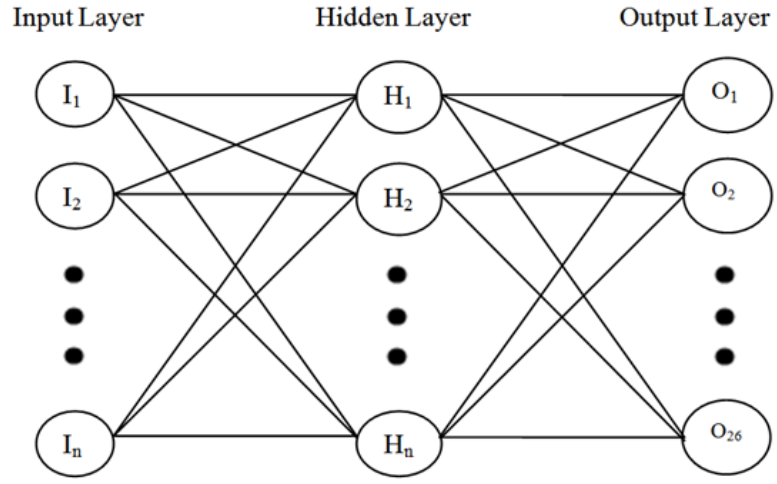

Figure 4. Architecture of Back Propagation Neural Network

In this experiment, the input features are based on the feature extraction methods. There are 16 features for zoning algorithm, 256 features for projection profile and 36 features for HOG feature extraction. The number of hidden nodes obtained from:

$$H = \frac{(x+y)}{2} + 1 \tag{5}$$

where H is the number of hidden nodes, x is the number of input features, and y is the total of classes, which is 26 classes in this experiment. Based on the formula to calculating the hidden nodes, it means that this experiment uses different number of hidden nodes. The following are the exact number of hidden nodes that used in this experiment: 22 hidden nodes for zoning algorithm, 142 hidden nodes for projection profile algorithm, and 32 hidden nodes for HOG algorithm. Some of the parameters that used in this experiment are, the training cycles are 500 and 1000, learning rate is 0.3, and momentum is 0.2 for every algorithm.

5. Experimental Results

This experiment uses the Chars74K dataset with the type of English Fnt set. Total images that used in this experiment is 26416 and limited to capitalized characters (A-Z) which each of characters contains 1016 different style of images. The composition of training and testing set is divided into 70:30. After that, the features of each image are extracted using different methods, which are zoning algorithm, projection profile, Histogram of Oriented Gradients (HOG), and the combinations of them (zoning + projection, projection + HOG, zoning + HOG, projection + zoning + HOG). There are two classifiers that used in this experiment: Support Vector Machine using RBF kernels, and the neural network. Table 1 shows the accuracy result using SVM while Table 2 shows the accuracy result using Neural Network.

Table 1. Accuracy Result using Neural Network

Feature Extraction	Accuracy Result 500 training cycles	Accuracy Result 1000 training cycles
Projection Profile	93.48%	93.48%
Zoning	86.07%	86.07%
HOG	93.48%	93.48%
Projection + Zoning	91.11%	91.11%
Projection + HOG	93.08%	93.08%
Zoning + HOG	86.10%	86.10%
Projection + Zoning + HOG	91.51%	91.51%

Table 2. Accuracy Result using Support Vector Machine

Feature Extraction	Accuracy Result
Projection Profile	**94.43%**
Zoning	81.21%
HOG	69.80%
Projection + Zoning	80.23%
Projection + HOG	**94.43%**
Zoning + HOG	81.21%
Projection + Zoning + HOG	80.23%

Overall, ANN resulted in higher accuracy compared to SVM, except for Projection and Projection + HOG feature extraction. Generally, using 500 training cycles in ANN methods provides better accuracy compared to more than 500 training cycles.

6. Conclusion

In this paper, we have compared several feature extraction methods and their combination, such as: zoning algorithm, projection profile, HOG, zoning + projection, projection + HOG, zoning + HOG, and zoning + projection + HOG. Those features were trained using 2 kinds of classifier, such as: SVM and ANN. The experimental results show the highest accuracy came from the utilization of projection profile algorithm and combination of projection and HOG using SVM classifier that achieves the 94.43% of accuracy, compared to the other algorithm.

There are several areas for future work. The results needs to be verified with complex dataset, subsequently to perform text detection from the hand writing result, low quality image, or from the video format that make the process of text detection more challenging. In addition to the results, the improvements of the proposed algorithm are needed.

References

1. Sharma S, Seth N. Survey of Text CAPTCHA Techniques and Attacks. Int J Eng Trends Technol. 2015;22(6).
2. Qadri MT, Asif M. Automatic Number Plate Recognition System for Vehicle Identification Using Optical Character Recognition. 2009 International Conference on Education Technology and Computer. 2009. p. 335–8.
3. Patel C, Shah D, Patel A. Automatic Number Plate Recognition System (ANPR): A Survey . 2013;69(9):21–33.
4. Ghai D, Jain N. Text Extraction from Document Images- A Review. 2013;84(3):40–8.
5. Malon C, Uchida S, Suzuki M. Mathematical symbol recognition with support vector machines. Pattern Recognit Lett [Internet]. 2008;29(9):1326–32. Available from: http://www.sciencedirect.com/science/article/pii/S0167865508000603
6. Technology AI, Rao NV, Pradesh A, Pradesh A, Pradesh A. OPTICAL CHARACTER RECOGNITION TECHNIQUE. 2016;83(2).
7. Mahto MK, Bhatia K, Sharma RK. Combined horizontal and vertical projection feature extraction technique for Gurmukhi handwritten character recognition. 2015 International Conference on Advances in Computer Engineering and Applications. 2015. p. 59–65.
8. Murthy OVR. Zoning based Devanagari Character Recognition. 2011;27(4):21–5.
9. Campos TE de, Babu BR, Varma M. Character recognition in natural images.
10. Impedovo S, Lucchese MG, Pirlo G. Optimal zoning design by genetic algorithms. Vol. 36, IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans. 2006. p. 833–46.
11. Zramdini A, Ingold R. Optical font recognition from projection profiles. 1998;6(SEPTEMBER 1993):249–60.
12. Bahi H El, Mahani Z, Zatni A, Saoud S. A robust system for printed and handwritten character recognition of images obtained by camera phone A robust system for printed and handwritten character recognition of images obtained by camera phone. 2015;(October).
13. Khorashadizadeh S, Latif A. Arabic/farsi handwritten digit recognition using histogram of oriented gradient and chain code histogram. Int Arab J Inf Technol. 2016;13(4):367–74.
14. Karatzoglou A, Meyer D, Hornik K. Journal of Statistical Software. 2006;15(9).
15. CHEY C, KUMHOM P, CHAMNONGTHAI K. KHMER PRINTED CHARACTER RECOGNITION BY USING WAVELET DESCRIPTORS. Int J Uncertainty, Fuzziness Knowledge-Based Syst [Internet]. 2006 Jun 1;14(3):337–50. Available from: https://doi.org/10.1142/S0218488506004047
16. Bhavsar H, Panchal MH. A Review on Support Vector Machine for Data Classification. 2012;1(10):185–9.
17. Jafri R, Arabnia HR. A Survey of Face Recognition Techniques. J Inf Process Syst. 2009;5(2):41–68.